Preprints.org

Article

Enhancing Legal Sentiment Analysis: A CNN-LSTM Document-Level Model

Bolanle Abimbola [*] , Enrique De La Cal Marin , Qing Tan [*]

Article

Enhancing Legal Sentiment Analysis: A CNN-LSTM Document-Level Model

Bolanle Abimbola [1],*, Enrique de La Cal Marin [1] and Qing Tan [2]

[1]   University of Oviedo; delacal@uniovi.es
[2]   Athabasca University, qingt@athabascau.ca
*   Correspondence: UO285018@uniovi.es

Abstract: This research investigates the application of deep learning in sentiment analysis of Canadian maritime case law. It offers a framework for improving maritime law and legal analytics policy-making procedures. The automation of legal document extraction takes center stage, underscoring the vital role sentiment analysis plays at the document level. Therefore, this study introduces a novel strategy for sentiment analysis in Canadian maritime case law, combining sentiment case law approaches with state-of-the-art deep learning techniques. The overarching goal is to systematically unearth hidden biases within case law and investigate their impact on legal outcomes. Employing Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) based models, this research achieves a remarkable accuracy of 98.05% for categorizing instances. In contrast, Conventional Machine Learning techniques such as Support Vector Machines (SVM) yield an accuracy rate of 52.57%, Naive Bayes at 57.44 %, and Logistic Regression at 61.86%. The superior accuracy of the CNN and STM model combination underscores its usefulness in legal sentiment analysis, offering promising future applications in diverse fields like legal analytics and policy design. These findings mark a significant choice for AI-powered legal tools, presenting more sophisticated and sentiment-aware options for the legal profession.

Keywords: Convolutional neural networks; deep neural networks; long short-term memory; sentimental analysis; recurrent neural networks

1. Introduction

In today's dynamic and interconnected world, the significance of information spans various critical domains, including legal, political, commercial, or individual perspectives and many more. Recognizing the pivotal role that opinions play in shaping decisions and influencing outcomes, there is a growing need for automated tools to analyze sentiments effectively. Regarding this case, sentiment analysis emerges as a significant participant. Sentiment mining, or Sentiment Analysis, is a comprehensive natural language processing approach that can identify and classify textual data's emotional tone and subjective content. People are beginning to communicate their thoughts more quickly and in a shorter time, making the manual processing of many viewpoints quite tricky. Therefore, sentiment analysis proved extremely useful in this field [1–3]. By employing the sentiment analysis technique, stakeholders can also navigate the intricate layers of precedents and decisions, enhancing their capacity for nuanced interpretation and contributing to more informed decision-making and policy formulation [2].

Recently, a significant amount of research has been conducted on opinion mining and sentiment analysis by applying machine learning and deep learning in various domains [4–6]. Opinion and sentiment analysis activities have been improved with the application of several neural networks, such as Convolutional Neural Networks (CNNs), GRU (Gated Recurrent Unit) or LSTM (Long Short Term Memory), and Recurrent Neural Networks (RNNs) [7]. Additionally, machine learning and deep learning models excel in analyzing short texts, leveraging abundant datasets from social networks to identify opinions quickly. However, tackling longer documents presents a more intricate challenge, given the higher word count and complex semantic links between sentences. Researchers

are increasingly invested in developing advanced analysis techniques to extract nuanced points of view on specific subjects from this substantial data mass. Navigating through the intricacies of longer documents, they aim to enhance sentiment analysis accuracy and gain deeper insights into complex topics, reflecting the evolving landscape of text analysis. From a legal perspective, there is a discernible trend towards integrating cutting-edge technologies such as machine learning and sentiment analysis to enhance the analytical capabilities of legal practitioners. Rhanoui, et al. [8] utilized the CNN-BiLSTM model to analyze press articles and blog posts and reported almost 90.66% accuracy. Similarly, Tripathy, et al. [9] a hybrid machine model was employed to classify document-level sentiment and claim positive feedback. Hence, this technological infusion holds particular promise in Canadian Maritime Case Law, where the complexities of legal texts and the need for precise forecasting of court decisions pose significant challenges [10].

Thus, this study aims to use a novel combination of deep learning techniques, namely Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, to glean emotional insights from Canadian maritime case law papers. This technique fills a gap in the literature by providing a novel strategy for using sentiment analysis in the law, and it does so by focusing on the application of deep learning to the study of Canadian maritime case law [13]. Legal sentiment analysis has the potential to completely alter how lawyers and judges examine massive collections of case law. The document's emotional tone, judgments, and sentiment dynamics are insightful for attorneys, judges, politicians, and academics. This study introduces deep learning models for examining Canadian maritime case law, which may pick up on subtleties of feeling that more conventional approaches would otherwise miss. Explored are the potential effects of these cutting-edge computational methods on legal analytics, policy formation, and the creation of AI-powered legal instruments.

The process begins with presenting case law, followed by an emotional evaluation of the findings. An extensive literature review explores the topic of sentimental analysis and its relevance to the legal profession. Then, the process involves gathering data to develop an ML model and analyze experiment outcomes aligning with prior research[14].

2. Background and Context: Sentiment Analysis

2.1. Sentiment Analysis

Sentiment analysis (SA) is a technological evaluation of people's thoughts, attitudes, and feelings toward a given object, which can be positive, negative, or neutral. Therefore, in this research, deep learning methods are used to solve the problem of extracting sentiment insights from Canadian maritime case law texts. Through the adept training of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks (see Figure 1), this novel approach opens new avenues for comprehending established legal doctrine and case law, facilitating better legal analysis and judgment. The implications extend far beyond the legal field, as deep learning models and algorithms can modernize the analysis of massive legal documents by revealing hidden emotions and how they impact the final judgment. With more significant implications in areas like legal analytics, policy design, and AI-powered legal tools, this study can potentially shape a more nuanced and well-informed legal environment [15].
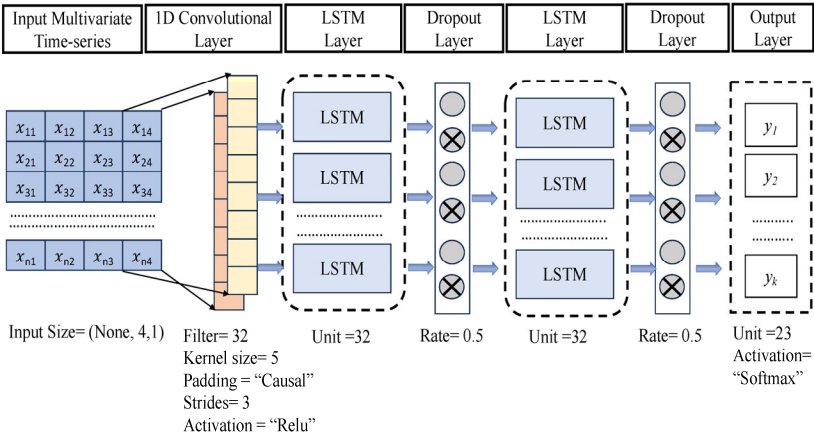
Figure 1. Canadian maritime case law texts by training Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks.

2.1.1. Levels of Sentiment Analysis

Opinion analysis involves sentiment analysis, assessing sentiments at both document and sentence levels. Sentiment analysis gauges the overall tone of a text and the sentence levels, providing broad and detailed insights, respectively. The aspect-based analysis focuses on specific elements or features, uncovering positive or negative feedback. Concurrently, sentence-level analysis is crucial in detecting and evaluating views directed at particular entities, offering a more intricate understanding of sentiments expressed within the text.

Comparative analysis involves assessing multiple entities or characteristics to ascertain their respective influences. Temporal analysis explores opinion evolution, trends, and the repercussions of events over time. In contrast, multilingual analysis uses various data formats, including text, images, audio, and video, to examine different points of view across multiple languages. Contextual analysis considers the contextual nuances that may alter words' meaning. The extent of opinion analysis depends on a particular application's specific objectives and conditions, which dictate the degree of opinion analysis used.

2.1.2. Word Embedding

Word embedding in natural language processing (NLP) is a remarkable strategy for enhancing sentiment analysis. Effectively unraveling the sentiments, attitudes, and views articulated in legal documents demands the deployment of sentiment analysis to discern the emotional tone of a text. The unparalleled ability of deep learning models, notably CNNs and LSTMs networks, to detect intricate patterns within text data has positioned them as the gold standard for sentiment analysis. Word embedding methods like Word2Vec, GloVe, and FastText play a critical role by mapping specialized legal lexicons into numerical vectors to transmute the semantic richness of words while translating them into numerical representations. Deep learning algorithms then harness these embeddings to decode feelings, proficiently capturing emotionally charged phrases and the nuanced deployment of language within context. The seamless integration of word embedding, and deep learning methodologies is indispensable in advancing sentiment analysis within Canadian Maritime Case Law. The consequential insights from this amalgamation hold immense value for legal practitioners, policymakers, and researchers, furnishing a nuanced comprehension of the emotional dimensions embedded in legal discourse [16].

2.2. Deep Learning

Deep learning, a subfield of machine learning and artificial intelligence (see Figure 2), focuses on training artificial neural networks to excel in tasks like data processing, pattern recognition, and decision-making. These networks, mirroring the structure and function of the human brain, consist

of interconnected layers of artificial neurons. One of the remarkable strengths of these models is their ability to automatically extract features and patterns, rendering them invaluable for applications such as sentiment analysis. Noteworthy designs within the realm of deep learning include CNNs, RNNs, and LSTMs. Particularly adept at deciphering intricate patterns and uncovering interdependencies in data, these networks find their effectiveness amplified in domains marked by complex terminology and nuanced relationships.
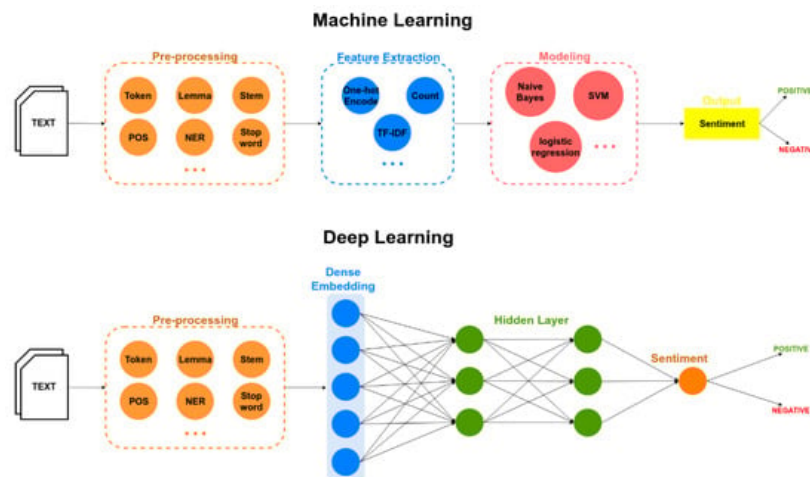

Figure 2. Sentiment polarity categorization using machine learning (above) and deep learning (below).

The following are some of the sentiment analysis aspects that deep learning models can handle:

Feature Extraction Word-to-word associations, the sentiments conveyed by individual words, and the overall context are all things deep learning models can deduce automatically.

Context Understanding is a comprehensive capability of capturing the contextual details essential for gaining a sense of emotion in complicated fields such as maritime law.

Sequential Information Modelling can efficiently generate sequential information, like RNNs and LSTMs models, which is essential for tasks requiring text order and sentiment. This is of utmost significance in legal documents, where structure and flow of information are critical.

Scalability complexity of Canadian Maritime Case Law is well within the capabilities of deep learning models. These models can handle vast datasets and be trained for specialized tasks.

While grappling with computational complexity and the necessity for fine-tuned hyperparameter adjustment, deep learning models are a remarkably effective tool in sentiment analysis within legal domains. As underscored in reference [17], these models extract subtle insights, offering a valuable conduit to elevate legal analytics. Their nuanced capabilities empower a more profound comprehension of sentiment within legal analytics and furnish a robust foundation for refining policy choices.

2.2.1. CNN

Convolutional Neural Networks (CNNs) are robust computational frameworks capable of decoding complex patterns within visual and textual data, particularly in sentiment analysis. These networks leverage spatial hierarchies to uncover subtle nuances in data, making them ideal choices for extracting important characteristics and patterns from textual data. CNNs apply filters or kernels to input data segments through strategic convolutional layers, precisely detecting sentiment-related words or phrases. This systematic feature extraction enables CNNs to uncover complex legal language.

A pooling layer in the convolutional layer efficiently reduces input dimensionality, capturing essential textual characteristics for sentiment analysis. By pinpointing critical sentiment-associated words, max-pooling selectively extracts pertinent information. CNNs leverage this interplay to create

sentiment parameters through iterative training on labeled text-sentiment data and bridge the gap between predictions and accurate sentiment labels. Pre-trained word embedding like Word2Vec and GloVe are employed to capture semantic relations to aid in interpreting legal contexts. For the best results when analyzing legal texts for sentiment, it's important to adjust hyperparameters such as CNN architecture, kernel size, filter number, and others to achieve optimum results [18].

2.2.2. RNN-LSTM

Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units are powerhouse tools for processing sequential data, shining in areas like language translation, speech recognition, and sentiment analysis. Their secret weapon? The ability to tackle the vanishing gradient problem, a hiccup that often plagues traditional neural networks, makes them particularly adept at parsing complex sequences found in texts, such as legal documents [19].

LSTMs excel at understanding the nuances of language thanks to their design that captures long-distance dependencies within text [20]. This capability is amplified by bidirectional LSTMs, which look at text from both directions, ensuring a robust context grasp for accurate sentiment detection. Leveraging pre-trained word embeddings, these networks dive deep into the specialized vocabulary crucial for tasks requiring a comprehensive text analysis.

Training these networks involves meticulously adjusting their architecture, including the layers and units specific to LSTMs, optimizing them for tasks that demand an understanding of extended sequences, which makes RNN-LSTMs particularly valuable for projects like sentiment analysis in Canadian Maritime Case Law, offering insights into the shifting tones within legal documents over time.

2.2.3. RNN-BiLSTM

Bidirectional Long Short-Term Memory (BiLSTM) networks significantly advance recurrent neural networks, particularly in processing natural languages [21]. Unlike traditional LSTMs, BiLSTM incorporates two hidden layers, enabling the simultaneous processing of data in both forward and backward directions. This BiLSTM approach enhances the network's ability to capture context and dependencies in sequential data. This architecture has demonstrated notable effectiveness in various natural language processing tasks, showcasing its prowess in sentiment analysis, named entity recognition, and machine translation. The bidirectional nature of BiLSTM allows it to capture nuanced patterns and relationships within language structures, making it a valuable tool in the ever-evolving landscape of deep learning applications for natural language understanding.

A powerful tool in Natural Language Processing (NLP), BiLSTM combines the advantages of LSTM with bidirectional processing [22]. Placing words in sentences within their historical and prospective contexts helps clarify their meanings. However, the BiLSTM network has many potential uses, including machine translation, text categorization, and named entity identification. Integrating it into advanced designs like BERT achieves benchmark performance on NLP that is second to none. On the other hand, longer sequences provide challenges for it because of the amount of computational work involved. Architectures based on transformers, such as BERT and GPT, are preferred in natural language processing owing to their parallelism and scalability [23].

3. Related Works

3.1. Short Text Sentiment Analysis

Understanding the emotions conveyed in 140-character posts like tweets, product reviews, comments, and status updates is the goal of the specialized discipline of Short Text Sentiment Analysis. Since more and more of our digital communication consists of concise sentences, this area has attracted a lot of study. While lengthier papers with more context may be analyzed using standard sentiment analysis techniques, brief texts with condensed and constrained characters introduce new obstacles. Extracting and analyzing feelings from brief writings, particularly in social

media and online reviews, is crucial since they provide vital information about the author's emotional tone and viewpoints.

Difficulties arise when analyzing the tone of short texts due to factors such as the absence of context, the use of informal language, background noise and abbreviations, and an unequal distribution of social classes. Since words and phrases in text messages may have multiple meanings, which can change depending on the context in which they are used, context is essential for deciphering emotions. Slang, conversational phrases, and emoticons/emojis provide unique challenges to emotion analysis because of their informal nature. Since noise and abbreviations might impact sentiment analysis findings, they are not ideal for brief text conversations [11].

Sentiment analysis models may be biased if there is a significant racial or ethnic minority in the population. Emoticon and emoji analysis, deep learning models like recurrent neural networks and convolutional neural networks, and transfer learning approaches are just a few specialized methods researchers and data scientists created for brief text sentiment analysis. Public opinion, consumer satisfaction, and new trends may all be gauged with these methods, which can be utilized in various contexts such as social media monitoring, customer service, and legislation.

3.2. Document Level Sentiment Analysis

Document-level sentiment analysis focuses on analyzing sentiment in lengthy texts such as articles, reviews, and reports, providing deep insight into emotional nuances and context. Unlike short-text analysis, which deals with concise texts, document-level analysis benefits from more comprehensive information to understand complex emotions in large texts, which is crucial for detailed sentiment comprehension applications [24]. It incorporates aspect-based evaluation for in-depth opinions on specific topics, necessitating algorithms capable of effectively handling sarcasm, ambiguity, and complex expressions. Machine learning models, including SVM, Naive Bayes, and RNNs, excel in this domain by capturing contextual nuances, which is vital for dissecting sentiments in product reviews and other detailed documents. Document-level analysis plays a significant role in natural language processing, supporting decision-making across various sectors by analyzing sentiments in product evaluations, financial reports, and social media [25]. Legal texts assist in identifying positive or negative sentiments, with neural networks offering advantages over traditional models by eliminating the need for explicit feature definitions [26]. Integrating AI and machine learning in law transforms traditional practices, enhancing document analysis and prediction accuracy. However, with the growing integration of AI in legal processes, ethical considerations and potential biases require careful management, particularly concerning data privacy and the ethical use of AI in judicial decisions [20]. Here, Table 1 provides an overview of the related work.

Table 1. Related Works.

Word Embedding	Level	Model	Accuracy
WORD2VEC	Word level	CNN-	84.9%
	Document	LSTM	84.7%
	level	BERT	89.0%
	Sentence	KNN	85.01%
	level	SSR	
GLOVE	Document	CNN-	88.9%
	level	BiLSTM	82.7%
	Word level	KNN	81.0%
	Sentence	CNN	91.01%
	level		
BOMW	Sentence	BOMW	92.9%
	level	BERT	78.7%
	Word level	CNN	86.0%

| Document level | SR-LSTM | 80.01% |

4. Proposed Model: CNN - LSTM and Doc2vec for Document-Level Sentiment Analysis

Cutting-edge methods in document-level sentiment analysis, like CNN-LSTM and Doc2Vec (see Figure 3), leverage advanced techniques to extract valuable insights and sentiment information from extensive texts like reviews, articles, and reports. These methods aim to decipher the text's underlying meaning and emotional nuances by employing deep learning and vector representations.

Although more commonly associated with image processing, CNNs can also be effectively trained for text analysis. In this context, CNNs are crucial in mining textual data to discern sentiment at the document level. Their functionality involves passing the input text through convolutional filters, adept at identifying local patterns and characteristics that may serve as sentiment indicators. The strength of CNN lies in its ability to pinpoint key phrases or sequences of words inside texts, thereby contributing significantly to the overall sentiment of the document.
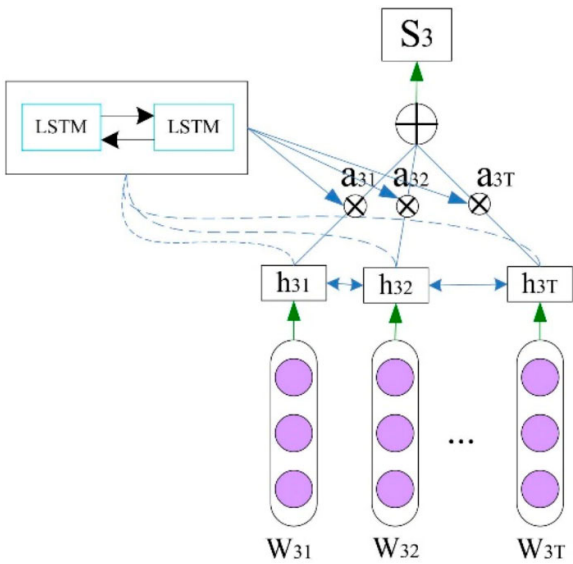

Figure 3. CNN - LSTM and Doc2vec for Document-Level Sentiment Analysis.

Long short-term memory (LSTM) RNNs excel in modeling sequential data, making them highly effective for understanding text's natural flow and context. In document-level sentiment analysis, LSTMs are crucial for extracting word and sentence dependencies, allowing them to capture evolving attitudes throughout lengthy texts. Their ability to selectively retain and forget information over extended sequences ensures consistent and nuanced sentiment analysis, making LSTMs indispensable in natural language processing.

Gates:

LSTMs use three types of gates: i) the forget gate (f), ii) the input gate (i), and iii) the output gate (o).

These gates control the flow of information into and out of the cell state (C_t).

a. Cell State (C_t):

The cell state represents the memory of the LSTM. It can be updated and modified using the gates.

The cell state is updated using the forget gate, input gate, and a new candidate cell state($C{\sim}\_t$).

b. Hidden State(h_t):

The hidden state carries information about the current time step's input and the previous hidden state.

It is used to make predictions and updated using the output gate.

Forget Gate(f_t):

$$f\_t \ = \ \sigma(W\_f \ \cdot \ [h\_(t-1), x\_t] \ + \ b\_f)$$

Here, σ represents the sigmoid activation function.

c. Input Gate (i_t):

$$i\_t \ = \ \sigma(W\_i \ \cdot \ [h\_(t-1), x\_t] \ + \ b\_i)$$

d. Candidate Cell State ($C\sim\_t$):

$$C\sim t \ = \ tanh(W\_c \ \cdot \ [h(t-1), x\_t] \ + \ b\_c)$$

e. Update Cell State (C_t):

$$C\_t \ = \ f\_t * C\_(t-1) \ + \ i\_t * C\sim\_t$$

This equation combines the old and new candidate cells based on forget and input gates.

f. Output Gate (o_t):

$$o\_t \ = \ \sigma(W\_o \ \cdot \ [h\_(t-1), x\_t] \ + \ b\_o)$$

g. Hidden State (h_t):

$$h\_t \ = \ o\_t * tanh(C\_t)$$

The output gate controls the information that is passed to the hidden state. However, here x_t represents the input at time step t, $h\_t - 1$ presents the hidden state at time step t-1. Similarly, W_f, W_i, W_c, W_o and b_f, b_i, b_c, b_o represent Weight matrices for the gates and Bias vectors for the gates respectively. On the other hand, σ stands for the sigmoid activation function, whereas $tanh$ presents the hyperbolic tangent activation function.

In deep learning, LSTM models are pivotal for language modeling, translate texts, and speech recognition due to their selective updating and retrieval processes. Combining convolutional neural network (CNN) with long short-term memory (LSTM) models proves a powerful approach for sentiment analysis, offering nuanced evaluations of entire articles, including positivity, negativity, or neutrality. In contrast to Word2Vec, Doc2Vec takes a holistic approach, representing entire manuscripts. This allows it to recognize emotional context and sentiments, even without labels or phrases, by analyzing papers comprehensively and understanding intricate linkages between words.

4.1. Model Overview and Motivation

Providing a comprehensive overview of the model and articulating the reasons behind its development are imperative steps in establishing the study's context, substantiating its significance, and justifying its importance.

The sentiment analysis model employed or under construction involves a sophisticated architecture comprising key components and methodologies pivotal to its functionality. The utilization of advanced techniques such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, or any other deep learning approaches. The utilization of advanced techniques, such as model construction, underscores its design's intricacies. An exhaustive elucidation of the model's development major components is paramount, necessitating a detailed exploration from the initial data preparation phase to the subsequent training and testing stages [27]. This comprehensive breakdown ensures a thorough understanding of the model's structure and the rationale behind the strategic integration of specific neural network architectures.

In delving into sentiment analysis within the context of Canadian marine case law, this model seeks to illuminate the distinctive features of this legal domain, emphasizing its uniqueness compared to broader legal corpora. The motivation for this specialized approach stems from the discernible gaps in the existing literature, where the intricacies of maritime law often remain underserved. By addressing these voids, the model aims to introduce fresh perspectives and techniques, showcasing innovative approaches to sentiment analysis. The practical applications of this model in real-world scenarios promise to enhance legal research, decision-making, and policy formulation within Canadian maritime case law. Furthermore, the article meticulously navigates the challenges and complexities inherent in sentiment analysis of legal documents, shedding light on the

shortcomings of current methodologies. Through its straightforward presentation of the model's structure and a compelling justification[28], this study not only fills an existing void in the understanding of sentiment in Canadian maritime case law but also has the potential to significantly influence the broader field of sentiment analysis within legal contexts.

4.2. Document Representation

Effective sentiment analysis in Canadian marine case law requires transforming legal texts into a suitable format for deep learning programs. The initial phase includes data cleansing, information removal, and standardization. Additionally, tokenization breaks down text into individual words, and word embedding represents them as numerical vectors. We use methods like averaging or TF-IDF weighting or more complex algorithms like Doc2Vec to produce document-level representations (see Figure 4).
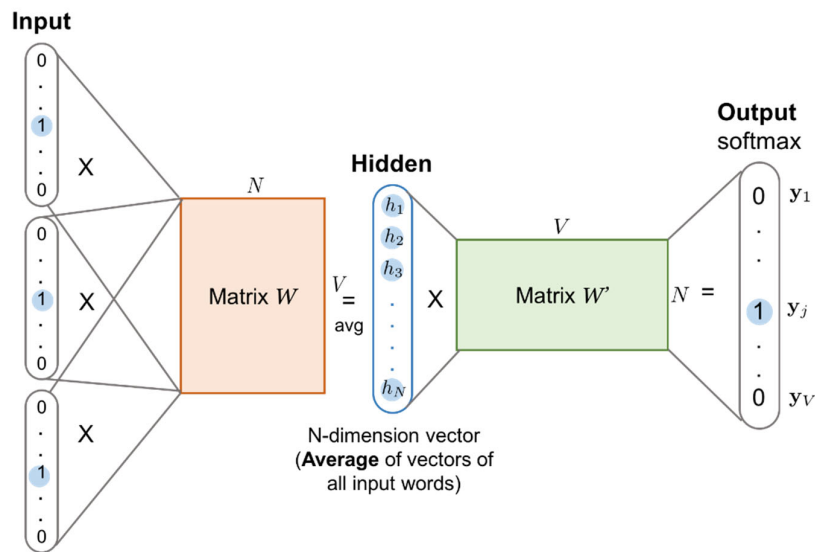

Figure 4. Word embeddings, which represent words as numerical vectors, are used to help deep-learning models comprehend the text.

In Canada's maritime case legislation, sequence padding is vital for efficiently processing legal documents with varying durations. Despite challenges like intricate legalese and lengthy paperwork, sequence padding is indispensable for sentiment analysis. It helps represent legal documents, enabling deep learning to discern emotional patterns within the previously unstructured legal material.

$$p(w0|wr) = \frac{exp\,(v'_{w0T_{vwr}})}{\sum_{i=1}^{v} exp\,(v'_{wi}T^{vw1})}$$

Like CBOW, Skip-Gram predicts every other word as an output after receiving a word as input.

The Continuous Bag of Words (CBOW) algorithm is a word prediction method considering the context. This paradigm is highly effective because it uses few resources and can be as simple as a single word or a group of words. This method is based on computing the negative logarithmic probability of a word (wt) in relation to its context$(r\hat{})$.

Word2Vec (Skip-gram with Negative Sampling):

Skip-gram:

Maximize the probability of the context words given the target word.

Equation:

$$P(context\,|\,target)\ =\ \Pi\,P(w\_context\,|\,w\_target)$$

The loss function (negative log-likelihood):

$$Loss\ =\ -\sum[log(P(w\_context\,|\,w\_target))]$$

Negative sampling (approximation to improve efficiency):

Maximize the probability of the true context words and minimize the probability of randomly sampled "negative" context words.

$P(w\_context \mid w\_target) = \sigma(w\_context \cdot w\_target)$ for true context words

$P(w\_negative \mid w\_target) = \sigma(-w\_negative \cdot w\_target)$ for negative samples

The loss function:

$$Loss = -[log(P(w\_context \mid w\_target)) + \sum[log(P(w\_negative \mid w\_target))]$$

$$GloVe\ (Global\ Vectors\ for\ Word\ Representation)$$

Minimize the squared difference between word vectors and their dot products.

$$J(\theta) = \sum(w\_i, w\_j)\,[f(P(w\_i \mid w\_j))*(w\_i \cdot w\_j - log(X\_i, j))\mathbin{\char`\^}2]$$

where X_i, j is the co-occurrence count of words w_i and w_j in the corpus, f is a weighting function, θ represents the model parameters.

FastText:

FastText introduces subword information and computes word vectors based on subword embeddings.

The subword vectors are combined to represent a word.

The equation for the word vector of a word is based on the summation of its subword vectors.

4.3. Convolution Layer

In Canadian maritime case law sentiment analysis, the Conv1D layer is crucial in CNN+LSTM models. Renowned for extracting intricate sentiment patterns from complex legal texts by identifying local textual patterns and navigating the length and intricacy of documents. (see Figure 5). By spotlighting these linguistic nuances, conv1D filters enhance the model's capacity to distinguish between positive and negative sentiments [29].
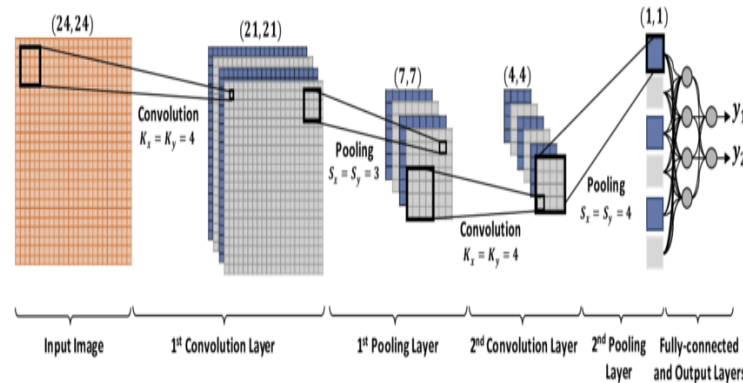

Figure 5. Input image processing.

"LeNet" and "AlexNet," two prominent CNNs, share linear neuron model principles. CNNs, unlike traditional MLPs, incorporate weight sharing and restricted connection in convolutional layers.

However, integrating Conv1D with LSTM layers proves a powerful tool for discerning emotions in complex legal documents. Conv1D's localized receptive model captures particular segments and global feelings, aided by max-pooling layers to preserve excessive feature loss. This approach benefits Canadian marine case law, providing a more accessible understanding of rulings and accommodating diverse perspectives.

The 1D forward propagation (1D-FP) expressions in each CNN layer are as follows:

$$x_k^l = b_k^l \sum_{i=1}^{N_{l-1}} conv1D\,(w_{ik}^{l-1}, s_i^{l-1})$$

The input is denoted by the symbol x_k^l, where b_k^l is the bias of the k th neuron at layer l, s_i^{l-1} is the output of the i th neuron at layer l-1, and w_{ik}^{l-1} is the kernel from the i th neuron at layer l-1 to the k th neuron at layer l.

$$y_k^l = f(x_k^l) \quad \text{and} \quad s_k^l = y_k^l \downarrow ss$$

With l = 1 as input and l as output, the back-propagation procedure begins at the MLP-layer. There are NL distinct types of data in the repository. In the output layer, we represent the mean-squared error (MSE) between an input vector p and its target and output vectors, t p and [y_k^l, yNL L]:

$$E_p = MSE\left(t^p, [y_l^L, \dots, y_{N_L}^L]'\right) = \sum_{i=1}^{N_L} (y_i^L - t_i^p)^2$$

E_p's derivative by each network parameter may be calculated using the delta error, k l = E xk l. To be more precise, the chain rule of derivatives may be used to update not just the bias of the current neuron but also the weights of all of the neurons in the layer above.
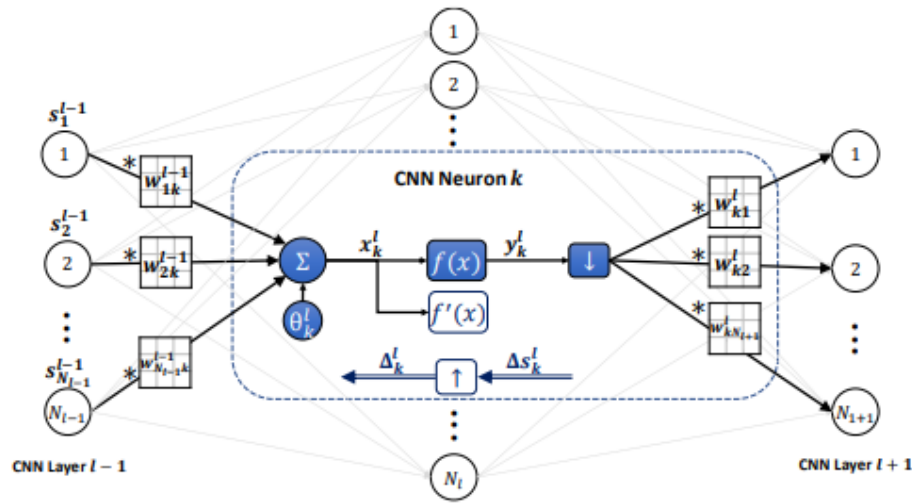

Figure 6. CNNs back-propagation and forward propagation.

CNNs with several layers use both back-propagation and forward propagation.
Through forward and reverse propagation, the last hidden CNN layer is linked to the first hidden MLP layer (see Figure 7)
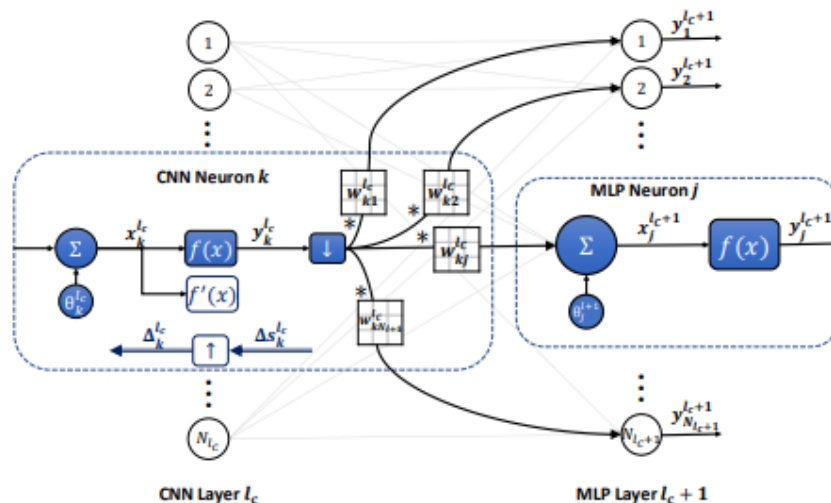

Figure 7. CNN layer linked to the first hidden MLP layer.

1) Initialize weights and biases (e.g., randomly, ~U(-0.1, 0.1)) of the network.
2) For each BP iteration DO:

    a.   For each PCG beat in the dataset, DO:

FP: A layer's neuron outputs may be found by forward propagation from the input layer to the output layer.

$$si\ l\ , \forall i\ \in\ [1, Nl\ ], and\ \forall l\ \in\ [1, L].$$

BP: Compute delta error at the output layer and back-propagate it to first hidden layer to compute the delta errors. $\Delta k\ l\ , \forall k\ \in\ [1, Nl\ ], and\ \forall l\ \in\ [1, L].$

     PP: Post-process to compute the weight and bias.

Update: Update the weights and biases by the (accumulation of) sensitivities scaled with the learning factor.

4.4. Activation Layer

In the context of sentiment analysis applied to Canadian marine case law texts using a CNN+LSTM architecture, the activation layer, also known as the activation function, is a pivotal element. The model's ability to capture complicated connections and produce precise predictions is greatly enhanced by adding this non-linear laNyer. By introducing non-linearity, the model becomes adept at discerning complex patterns, enabling more accurate predictions and nuanced insights into sentiment from the nuanced language of legal documents.

One of the most important parts of deep learning models is the activation layer, which helps to interpret complex legal documents' sentiment patterns and other nuanced emotional expressions in the data they include. This essential layer is the linchpin for capturing and learning from recurring structures, decision-making, and controlling gradient flow within the model. Despite its undeniable significance, the activation layer is not without challenges, with the specter of saturation looming as a potential impediment to the deep learning model's learning speed and overall effectiveness. Nevertheless, its indispensability remains unassailable, as the success of deep learning models in the nuanced domain of sentiment analysis within legal texts is intricately tied to the adept functioning of the activation layer, as underscored by empirical evidence [30].

4.5. Regularization

Combining deep learning methods, like CNN with LSTM models for sentiment analysis in Canadian marine case law, heavily employs regularization techniques to counteract overfitting. The complexity of legal language patterns makes accurate representation critical. An issue is overfitting, which occurs when a model performs exceptionally well on training data but poorly on new data. To achieve robust sentiment analysis in this field, regularization is essential for ensuring the model can handle both the complexities of the training data and the wide variety of legal text patterns.

4.6. Optimization

Sentiment analysis is crucial in Canadian maritime case law to understand the complex feelings expressed in legal documents. Combining CNN and LSTM models is one example of a complex model used to improve sentiment analysis. These models undergo optimization as part of the training process, which entails fine-tuning many parameters. The main goal is to improve the model's forecast accuracy by reducing its loss function. The optimization of sentiment analysis models becomes crucial in the complex world of legal discourse to guarantee a detailed comprehension of the emotions undercurrents in marine case law. The model's sentiment detection capabilities are enhanced, and the intricacies of legal language are more accurately reflected through this repeated tuning process.

4.7. BiLSTM Layer

Data is consistently organized sequentially in this layer, facilitating a structured representation. The layer explores connections between inputs and outcomes, while at its entry point, outcomes from maximum pooling operations are concatenated. This concatenation enhances the model's capacity to capture diverse patterns, promoting a robust understanding of input data and contributing to effective learning and predictions.

5. Experimental Results

5.1. Data Set

Legal documents in Canada are organized into several types, with Maritime Law legislation being just one example. Many techniques are employed for data classification, including text mining, document clustering, and machine learning algorithms. CNN is one of the models that is often used in the document classification process. The main tools were to predict a judge or jury's decisions and examine previous cases and decisions. However, in some instances, machine learning algorithms can make it easier to consider releasing a suspect on bail. In this research, two thousand cases were analyzed from the Federal High Court's website to identify patterns in maritime law in Canada (see Table 2). Additionally, the data was collected manually, without using anonymization, from both the plaintiffs and defendants.

Table 2. Features identified in the data.

Case Year	The year the case was registered.
Majority Opinion	Opinion of the majority of judges engaged in the case.
Minority Opinion	Opinion of the minority of judges engaged in the case.
Number of judges	The total number of judges hearing the case.
Court Judgment	Final court judgment on the case (whether the decision is affirmed or not).
Number of cited documents (Court decision legislation data)	The number of laws and judicial jurisprudence cited by the judges to support their decision.

To enhance sentiment analysis within maritime law, this research strategically uses the filter tool available on the Federal High Court website. This tool facilitated the identification of pertinent legislation and precedents from court rulings. A meticulous augmentation process was undertaken to bolster the sentiment analysis model, generating an additional 98,000 new samples through a random sample technique. This method deliberately addressed demographic disparities, ensuring a more even distribution of examples across emotion categories. The resultant effect was a marked improvement in the model's precision and consistency. Notably, the model's accuracy in categorizing emotions is fortified by incorporating Canadian marine case law. This deliberate and rigorous approach to data augmentation contributes significantly to the overall trustworthiness and precision of the sentiment analysis methodology employed in this study [31].

5.2. Results

The study on case adjudication in Canadian maritime law revealed intriguing insights into the outcomes of trials based on the number of judges involved. When a case was assigned to a single judge, guilty judgment stood at 46%, while the likelihood of a not guilty result was approximately 51%. Strikingly, this indicated a remarkably even distribution of judgments, with approximately 3% of cases remaining undecided. Surprisingly, the incidence of indecisive verdicts did not significantly change when three judges were involved, as it increased marginally to 5%. These findings suggest that the presence of additional judges in the trial process did not substantially alter the proportion of undecided cases, highlighting a noteworthy consistency in judgment outcomes across varying

judicial scenarios in the realm of Canadian maritime law. Surprisingly, the incidence of indecisive verdicts did not exhibit a significant shift when three judges.

Accuracy is the rate at which a model makes accurate predictions.

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$$

In Canadian maritime law, a significant shift has occurred in citation practices, revealing 41% of citations in single-judge trials and 46% in multi-judge cases. This evolving trend underscores the dynamic nature of the legal landscape. Scholars employ advanced techniques for sentiment analysis on Canadian marine case law papers, including deep learning and traditional machine learning models. This analytical approach extends beyond statistics, offering valuable insights for informed decision-making in judge selection and jury verdicts [32]. Integrating technology into legal scholarship reflects a proactive response to contemporary challenges, enhancing the adaptability of legal practices.

Figure 8 is a comprehensive visual representation of a bar chart, elucidating the distribution of judgments and statuses throughout the dataset. Each bar's height succinctly encapsulates the number of instances within its corresponding category, offering a clear and insightful overview of the dataset's composition. This visualization lays a robust foundation for forthcoming legal sentiment analysis studies and provides vital insights into the dataset's composition, knowing the predominance of judgments in Canadian marine case law [33].
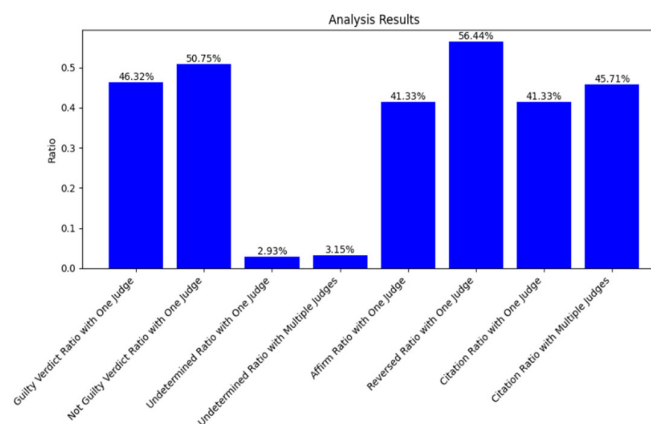

Figure 8. Dynamics of case adjudication.

In the initial stages of model assessment (see Figure 9), the dataset was carefully split into training and test sets, with non-predictive columns removed from the feature matrix X. The target variable y was appropriately labeled "target" for the subsequent binary classification task. To ensure repeatability, 30 percent of the data set was reserved for thorough examination. Preceding sentiment analysis, the 'Opinion' text input underwent tokenization to achieve consistent sequence lengths in the CNN with LSTM model. The vocabulary size was calculated, encompassing all distinct words in the 'Opinion' text data, and the longest sequence in the dataset (max_len) was identified. These preprocessing steps are vital for the success of sentiment analysis performed on Canadian marine case law materials that require this preliminary processing [33].
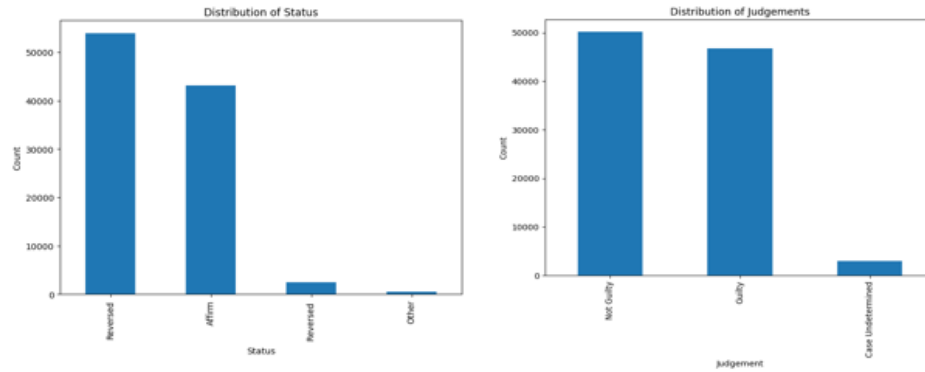
Figure 9. Distribution of status and judgments in the dataset.

5.3. Comparison

This section compares CNN, LSTM, BiLSTM, and CNN-LSTM to the CNN-BiLSTM model.

This study examines the sentiment analysis of Canadian maritime case law using two machine learning models: deep learning (CNN+LSTM) and more traditional methods (Logistic Regression, Multinomial Naive Bayes, Linear Support Vector Machine). The study's success is credited with employing CNN and LSTM models to collect sentiment information from judicial documents.

5.3.1. CNN Model

The ability of CNN models to extract local patterns and features from text input makes them particularly well-suited to tasks that require recognizing nearby signals or characteristics. They can spot terms, phrases, or clauses in legal papers that convey emotion. CNNs provide computational efficiency during training because of their ability to learn local patterns quickly through utilizing shared weights across several input areas. The training time is drastically reduced, making it particularly suitable for big legal text datasets. In this case, CNNs are powerful feature extractors that can glean important information from texts, including patterns, structures, or even individual words. The local environment brief pieces of text are quickly captured by them, and they excel at identifying patterns within such sections. The effectiveness of a CNN model in detecting emotions in Canadian maritime case law papers was demonstrated and found by its 98% accuracy rate in the tests [34].

5.3.2. LSTM Models

Long short-term memory (LSTM) models excel in understanding the context and sequence of words in text data, making them excellent for jobs requiring such an understanding. A thorough comprehension of the complex textual environment is essential in legal sentiment analysis. LSTMs are well-suited to the level of detail needed to comprehend the nuanced sentiment patterns and intricate interconnections common in legal writings. LSTM models are more complex and have a more significant number of parameters than CNN models. However, they still achieve high accuracy rates, which reflects how well they read sentiment dynamics in Canadian maritime case law. LSTM models are more complex and have a more significant number of parameters than CNNs. However, they still achieve high accuracy rates, which reflects how well they read sentiment dynamics in Canadian maritime case law.

5.3.3. CNN-LSTM Model

The study assesses the efficacy of CNN and LSTM models over 50 training epochs using visual representations of loss and accuracy measurements. While the accuracy graph illustrates how well the model can classify data, the loss graph shows how well it can reduce inaccurate predictions. In addition, the loss graph reflects the model's skill in minimizing prediction mistakes, whereas the accuracy graph depicts its skill in accurately incorporating labels into opinions. By making it more

straightforward to visualize how the model evolved during training to integrate documents from Canada's marine case law, the SE visuals add to the broader discussion on sentiment analysis in the law [35].

Each CNN+LSTM model displays loss and accuracy graphs across 50 iterations during training.

In a groundbreaking study of Canada's maritime sector, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models were employed to analyze case law and identify patterns of emotion. The impressive successes in emotion categorization, as depicted in Table 3, underscore the complexity of emotion in this intricate area of law. This research also showcases the effectiveness of advanced machine learning in navigating the challenging landscape of maritime law, where understanding and addressing emotions add an additional layer of complexity for legal professionals.

An impressive 98.01% test accuracy rate was achieved by CNN and LSTM Model 1 (see Figure 10), showcasing its dominance in sentiment categorization, and understanding of the intricacies of Canadian maritime case law texts. On the other hand, Model 2 (see Figure 11), a descendant of Model 1, highlighted the robustness of the CNN+LSTM architecture with a test accuracy of 97.94%, proving its efficacy in extracting sentiment information from dense legal texts.

Similarly, Model 3 (see Figure 12) achieved a test accuracy rate of 98.05%, and the third model earned a test accuracy rate of 98.05%, demonstrating the approach's resilience in predicting sentiment dynamics within Canadian maritime case law despite the continuously high accuracy rates of CNN and LSTM models.
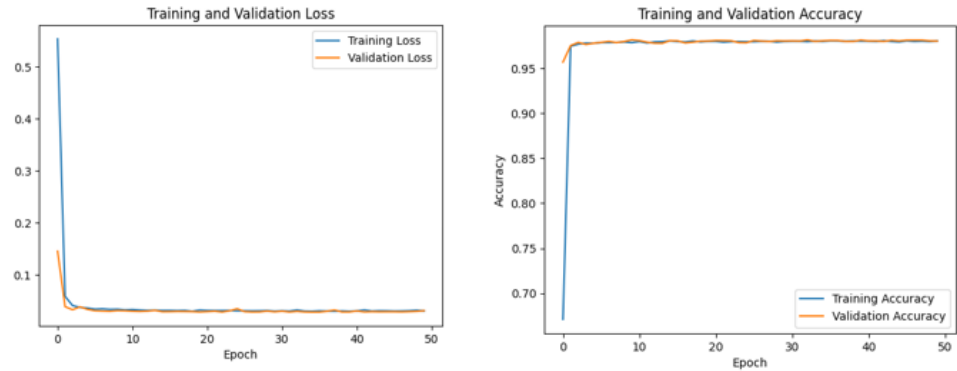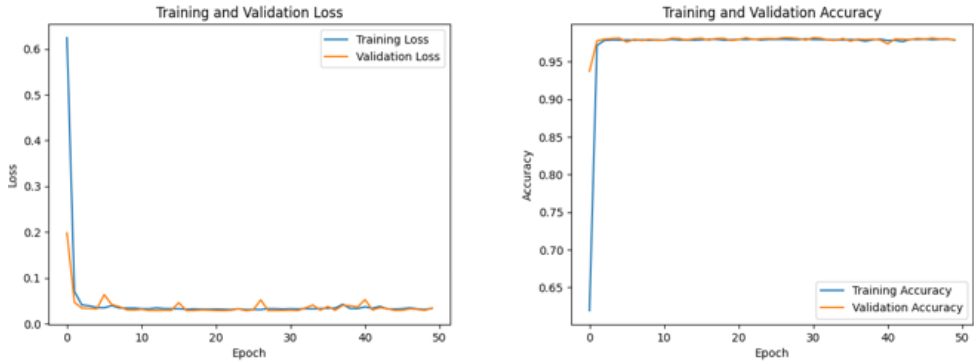

Figure 10. CNN+LSTM model 1.

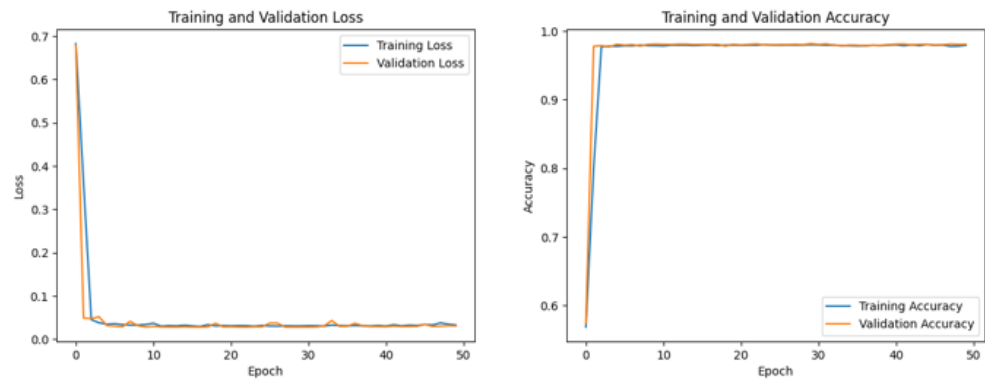**Figure 11.** CNN+LSTM model 2.

Figure 12. Loss and accuracy graphs of each CNN+LSTM model.

This research illustrates the effectiveness of CNN+LSTM models in analyzing legal sentiment analysis. It demonstrates how these models can more accurately detect sentiment patterns in Canadian maritime case law papers and successfully grasp the nuances of legal language. Legal analytics, policymaking, and the creation of AI-powered legal tools all stand to benefit significantly from this breakthrough [36].

Table 3. Results obtained for each deep learning model.

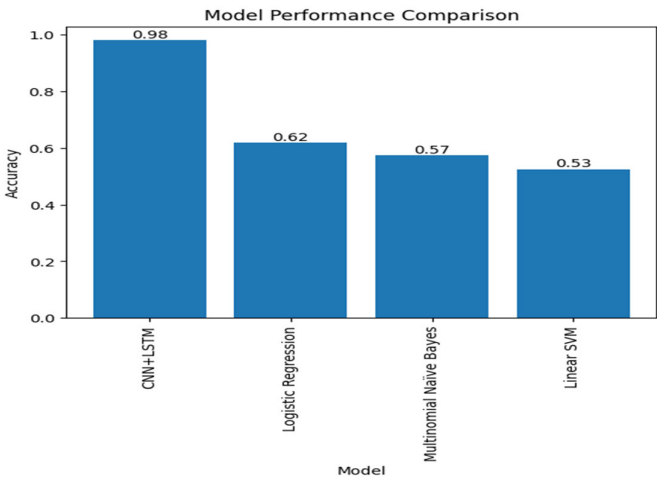| Model	Test Accuracy
CNN+LSTM model 1	98.01%
CNN+LSTM model 2	97.94%
CNN+LSTM model 3	98.05%

Figure 13. Model performance comparison for all models used.

Several machine learning algorithms were used to study Canadian maritime case law sentiment. With an average accuracy of 0.9805, CNN+LSTM models exhibited excellent precision in interpreting the nuances of legal documents. Logistic regression, multinomial naive Bayes, and linear support vector machine (SVM) are classic models that have significantly contributed to our knowledge of sentiment analysis by emphasizing the trade-offs between complexity, interpretability, and performance.

Multinomial Nave Bayes is practical with text data, whereas Logistic Regression sheds light on the effect of model complexity. In linear SVM, the emphasis is on parameterization and dataset dimensionality. This additional data will help us select more suitable policymaking and legal

analytics models. With this new information, we can better choose appropriate models for legal analytics and policy development [37].

5.4. Discussion

The CNN-BiLSTM with Doc2vec, a pre-trained sentence/paragraph representation model, stood out when we compared its performance to other deep learning models.

Doc2Vec word embedding models' accuracy ratings when using several neural network topologies, including convolutional neural networks (CNNs), long short-term memory (LSTM) networks, back-propagation neural networks (CNN-LSTM), and convolutional neural networks (CNN-BiLSTM), are offered. Document classification using Doc2Vec receives 90% on CNN, 88% on LSTM, 86.40% on BiLSTM, 91% on CNN-LSTM, and 93% on CNN-BiLSTM. Higher accuracy levels imply superior performance in sentiment analysis and text categorization [38].

This research explores the integration of deep learning, specifically Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models, in analyzing public opinion on maritime law in Canada. Achieving an accuracy above 98%, the study underscores the transformative impact of AI in the legal field, emphasizing automation of operations, extraction of meaning from extensive legal texts, and improved decision-making. The findings highlight both the benefits and drawbacks of these technologies, offering crucial insights for future applications. Significantly, sentiment analysis emerges as a valuable tool in various legal activities, including researching the law, investigating potential outcomes, preparing for court, interpreting precedent, and developing policies [39]. This research is a foundational step towards enhanced AI integration in legal practices, paving the way for further exploration and refinement in maritime law and beyond.

6. Conclusions

This study uses advanced deep learning methods, including convolutional neural networks (CNNs) and long short-term memory (LSTM) architectures, to unearth the nuanced feelings behind Canadian maritime case law. These results shed light on the subtleties of Canadian marine case law and the complex interplay between public opinion and judicial decisions. With an average accuracy of 98.05% across several examples, the CNN and LSTM models proved their ability to identify nuanced emotions in legal writing. This research shows that Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) are helpful for sentiment analysis in maritime law in Canada [40].

The models showed impressive accuracy ratings, with some reaching 98%. Convolutional neural networks (CNNs) effectively recognized local textual patterns, while long short-term memory (LSTM) models captured long-range relationships and sequential information. These models could provide valuable information for lawyers, enhancing investigations, evaluations, policymaking, legal analysis, and court strategy. AI-driven tools can provide fresh insights into complex issues and improve legal procedures [41].

The research highlighted the significance of parameter tuning and dataset dimensionality by comparing deep learning outcomes with traditional machine learning models and found that Logistic Regression achieved the highest accuracy (61.86%), Multinomial Nave Bayes showed 56.44% efficiency, and Linear Support Vector Machine depicted 52.5% efficient respectively, emphasizing the significance of parameter tuning and dataset dimensionality.

The research explores deep learning methods. Analyzing sentiment in maritime case law in Canada is the primary focus of this research, which focuses on applying deep learning techniques to legal analytics and policy creation. It highlights the importance of AI in legal practice and policy development and compares various machine learning models. The outcomes suggest that AI can produce a more complex and well-informed legal environment, claiming the potential of AI in legal practice [42].

Author Contributions: Conceptualization, B.A. and Q.T.; methodology, B.A.; validation, B.A., Q.T., and E.M.; formal analysis, B.A.; investigation, M.F.; resources, B.A. and Q.T.; writing—original draft preparation, B.A.;

References

1. B. Liu, *Sentiment analysis and opinion mining*. Springer Nature, 2022.
2. T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proceedings of the 2nd international conference on Knowledge capture*, 2003, pp. 70-77, doi: https://doi.org/10.1145/945645.945658.
3. X. Bai, "Predicting consumer sentiments from online text," *Decision Support Systems,* vol. 50, no. 4, pp. 732-742, 2011/03/01/ 2011, doi: https://doi.org/10.1016/j.dss.2010.08.024.
4. U. Naseem, I. Razzak, K. Musial, and M. Imran, "Transformer based Deep Intelligent Contextual Embedding for Twitter sentiment analysis," *Future Generation Computer Systems,* vol. 113, pp. 58-69, 2020/12/01/ 2020, doi: https://doi.org/10.1016/j.future.2020.06.050.
5. N. N. Yusof, A. Mohamed, and S. Abdul-Rahman, "Context Enrichment Model Based Framework for Sentiment Analysis," in *Soft Computing in Data Science: 5th International Conference, SCDS 2019, Iizuka, Japan, August 28–29, 2019, Proceedings 5*, 2019: Springer, pp. 325-335.
6. P. Vijayaragavan, R. Ponnusamy, and M. Aramudhan, "An optimal support vector machine based classification model for sentimental analysis of online product reviews," *Future Generation Computer Systems,* vol. 111, pp. 234-240, 2020/10/01/ 2020, doi: https://doi.org/10.1016/j.future.2020.04.046.
7. T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, "Recurrent neural network based language model," in *Interspeech*, 2010, vol. 2, no. 3: Makuhari, pp. 1045-1048.
8. M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM model for document-level sentiment analysis," *Machine Learning and Knowledge Extraction,* vol. 1, no. 3, pp. 832-847, 2019.
9. A. Tripathy, A. Anand, and S. K. Rath, "Document-level sentiment classification using hybrid machine learning approach," *Knowledge and Information Systems,* vol. 53, pp. 805-831, 2017.
10. K. NEWMYER and M. Zaccagnino, "VOLUME 52 FEBRUARY 2021 NUMBER 4," 2021.
11. R. F. Southcott and K. A. Walsh, "Canadian maritime law update: 2006," *J. Mar. L. & Com.,* vol. 38, p. 335, 2007.
12. O. M. Brandes and D. Curran, "Changing currents: A case study in the evolution of water law in Western Canada," *Water policy and governance in Canada,* pp. 45-67, 2017.
13. A. Christodoulou and J. Echebarria Fernández, "Maritime Governance and International Maritime Organization instruments focused on sustainability in the light of United Nations' sustainable development goals," in *Sustainability in the Maritime Domain: Towards Ocean Governance and Beyond*: Springer, 2021, pp. 415-461.
14. V. Gavrilov, R. Dremliuga, and R. Nurimbetov, "Article 234 of the 1982 United Nations Convention on the law of the sea and reduction of ice cover in the Arctic Ocean," *Marine Policy,* vol. 106, p. 103518, 2019/08/01/ 2019, doi: https://doi.org/10.1016/j.marpol.2019.103518.
15. S. Undavia, A. Meyers, and J. E. Ortega, "A comparative study of classifying legal documents with neural networks," in *2018 Federated conference on computer science and information systems (FedCSIS)*, 2018: IEEE, pp. 515-522.
16. B. Abimbola, Q. Tan, and J. R. Villar, "Introducing Intelligence to the Semantic Analysis of Canadian Maritime Case Law: Case Based Reasoning Approach," in *International Workshop on Soft Computing Models in Industrial and Environmental Applications*, 2022: Springer, pp. 587-595.
17. M. Ghorbani, M. Bahaghighat, Q. Xin, and F. Özen, "ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing," *Journal of Cloud Computing,* vol. 9, no. 1, pp. 1-12, 2020.
18. Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Computing and Applications,* vol. 32, pp. 9713-9729, 2020.
19. F. Gers, "Long short-term memory in recurrent neural networks," Verlag nicht ermittelbar, 2001.
20. S. Sohangir, D. Wang, A. Pomeranets, and T. M. Khoshgoftaar, "Big Data: Deep Learning for financial sentiment analysis," *Journal of Big Data,* vol. 5, no. 1, pp. 1-25, 2018.

21. M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing,* vol. 45, no. 11, pp. 2673-2681, 1997.
22. K. S. Tai, R. Socher, and C. D. Manning, "Improved semantic representations from tree-structured long short-term memory networks," *arXiv preprint arXiv:1503.00075,* 2015.
23. A. Sadia, F. Khan, and F. Bashir, "An overview of lexicon-based approach for sentiment analysis," in *2018 3rd International Electrical Engineering Conference (IEEC 2018)*, 2018, pp. 1-6.
24. S. Guo and G. Zhang, "Using machine learning for analyzing sentiment orientations toward eight countries," *Sage Open,* vol. 10, no. 3, p. 2158244020951268, 2020, doi: https://doi.org/10.1177/2158244020951268.
25. E. C. ATEŞ, G. E. BOSTANCI, and M. Serdar, "Big data, data mining, machine learning, and deep learning concepts in crime data," *Journal of Penal Law and Criminology,* vol. 8, no. 2, pp. 293-319, 2020, doi: https://doi.org/10.26650/JPLC2020-813328.
26. A. Kaur and B. Bozic, "Convolutional Neural Network-based Automatic Prediction of Judgments of the European Court of Human Rights," in *AICS*, 2019, pp. 458-469.
27. X. Li, X. Kang, C. Wang, L. Dong, H. Yao, and S. Li, "A neural-network-based model of charge prediction via the judicial interpretation of crimes," *IEEE Access,* vol. 8, pp. 101569-101579, 2020.
28. M. Medvedeva, M. Vols, and M. Wieling, "Using machine learning to predict decisions of the European Court of Human Rights," *Artificial Intelligence and Law,* vol. 28, pp. 237-266, 2020.
29. K. Machová, M. Mikula, X. Gao, and M. Mach, "Lexicon-based sentiment analysis using the particle swarm optimization," *Electronics,* vol. 9, no. 8, p. 1317, 2020, doi: 10.3390/electronics9081317
30. D. Alghazzawi, O. Bamasag, A. Albeshri, I. Sana, H. Ullah, and M. Z. Asghar, "Efficient prediction of court judgments using an LSTM+ CNN neural network model with an optimal feature set," *Mathematics,* vol. 10, no. 5, p. 683, 2022, doi: 10.3390/math10050683
31. A. Bramantoro and I. Virdyna, "Classification of divorce causes during the COVID-19 pandemic using convolutional neural networks," *PeerJ Computer Science,* vol. 8, p. e998, 2022.
32. J. Watson, G. Aglionby, and S. March, "Using machine learning to create a repository of judgments concerning a new practice area: a case study in animal protection law," *Artificial Intelligence and Law,* vol. 31, no. 2, pp. 293-324, 2023.
33. N. C. Da Silva *et al.*, "Document type classification for Brazil's supreme court using a convolutional neural network," in *10th International Conference on Forensic Computer Science and Cyber Law (ICoFCS), Sao Paulo, Brazil*, 2018, pp. 29-30.
34. V. G. Pillai and L. R. Chandran, "Verdict prediction for indian courts using bag of words and convolutional neural network," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020: IEEE, pp. 676-683.
35. D. L. Chen and J. Eagel, "Can machine learning help predict the outcome of asylum adjudications?," in *Proceedings of the 16th edition of the International Conference on Articial Intelligence and Law*, 2017, pp. 237-240.
36. K. Lum, "Limitations of mitigating judicial bias with machine learning," *Nature Human Behaviour,* vol. 1, no. 7, p. 0141, 2017.
37. A. Tasdelen and B. Sen, "A hybrid CNN-LSTM model for pre-miRNA classification," *Scientific reports,* vol. 11, no. 1, p. 14125, 2021.
38. F. Muhlenbach, L. N. Phuoc, and I. Sayn, "Predicting Court Decisions for Alimony: Avoiding Extra-legal Factors in Decision made by Judges and Not Understandable AI Models," *arXiv preprint arXiv:2007.04824,* 2020.
39. A. Alsayat, "Improving sentiment analysis for social media applications using an ensemble deep learning language model," *Arabian Journal for Science and Engineering,* vol. 47, no. 2, pp. 2499-2511, 2022.
40. J. T. Lam, D. Liang, S. Dahan, and F. H. Zulkernine, "The Gap between Deep Learning and Law: Predicting Employment Notice," in *NLLP@ KDD*, 2020, pp. 52-56.
41. B. Abimbola, "Sentiment Analysis of Canadian Maritime Case Law: A Sentiment Case Law and Deep Learning Approach " *Mendeley Data,* vol. V1, 2023, doi: doi: 10.17632/7v4jmbjwvc.1.
42. M. E. Alzahrani, T. H. Aldhyani, S. N. Alsubari, M. M. Althobaiti, and A. Fahad, "Developing an intelligent system with deep learning algorithms for sentiment analysis of E-commerce product reviews," *Computational Intelligence and Neuroscience,* vol. 2022, 2022, doi: https://doi.org/10.1155/2022/3840071.